

Speech recognition device for integrated services communications systems, in particular tel phone switching systems

Patent Number: DE3643838
Publication date: 1988-06-30
Inventor(s): ZINKE JOACHIM DIPL ING (DE)
Applicant(s): TELEFONBAU & NORMALZEIT GMBH (DE)
Requested Patent: ☐ DE3643838
Application Number: DE19863643838 19861220
Priority Number(s): DE19863643838 19861220
IPC Classification: G10L3/00 ; G06F3/16
EC Classification: H04Q11/04S, G10L15/00
Equivalents:

Abstract

The reference pattern memory provided for speech recognition is divided into a plurality of areas which are in each case assigned to a calling connection device. These areas are activated directly by the telephone number and/or address of a calling connection device which is transmitted via a signalling channel. Only the reference patterns of a single speaker or a small number of speakers which are stored in the area addressed in this way are fed to a comparator. A changeover device is provided which can be set by a control device in such a way that the speech information arriving via a user data channel is either forwarded to the input of the reference pattern memory for storage or is fed to the comparator for command generation.

Data supplied from the esp@cenet database - I2

THIS PAGE BLANK (USPTO)

①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ **Offenlegungsschrift**
⑪ **DE 3643838 A1**

⑤1 Int. Cl. 4:
G 10 L 3/00
G 06 F 3/16

②1 Aktenzeichen: P 36 43 838.3
②2 Anmeldetag: 20. 12. 86
④3 Offenlegungstag: 30. 6. 88

Behörden:igentum

DE 3643838 A1

⑦1 Anmelder:

Telenorma Telefonbau und Normalzeit GmbH, 6000
Frankfurt, DE

⑦2 Erfinder:

Zinke, Joachim, Dipl.-Ing., 6000 Frankfurt, DE

⑤4 **Spracherkennungseinrichtung für dienste-integrierende Kommunikationsanlagen, insbesondere Fernsprechvermittlungsanlagen**

Der für eine Spracherkennung vorgesehene Referenzmusterspeicher ist in mehrere jeweils einem anrufenden Anschlußorgan zugeordnete Bereiche unterteilt. Diese werden direkt durch die über einen Signalisierungskanal übertragene Rufnummer und/oder Adresse eines anrufenden Anschlußorgans angesteuert. Dabei werden nur die in dem so adressierten Bereich eingespeicherten Referenzmuster eines einzigen oder auch einiger wenigen Sprecher einem Vergleich angeboten. Es ist eine Umschalteneinrichtung vorgesehen, die von einer Steuereinrichtung so einstellbar ist, daß die über einen Nutzdatenkanal ankommende Sprachinformation entweder zum Einspeichern an den Eingang des Referenzmusterspeichers gelangt oder zur Be-
fehlserzeugung dem Vergleich angeboten wird.

DE 3643838 A1

EL302703433

Patentansprüche

1. Spracherkennungseinrichtung für dienste-integrierende Kommunikationsanlagen, insbesondere Fernsprechvermittlungsanlagen, wobei an zentraler Stelle ein Referenzmusterspeicher angeordnet ist, der Referenzmuster von sprachlich eingegebenen Befehlsworten verschiedener Sprecher aufnimmt, und ein Vergleichervorgesehen ist, der ankommende Sprachsignale mit den gespeicherten Referenzmustern vergleicht, um Steuerbefehle anhand der erkannten Befehlsworte zu erzeugen, dadurch gekennzeichnet, daß der Referenzmusterspeicher (RMS) in mehrere jeweils einem anrufenden Anschlußorgan (AO) zugeordnete Bereiche (TB 1 bis TBn) unterteilt ist, welche direkt durch die über einen Signalisierungskanal (D-Kanal) übertragene Rufnummer und/oder Adresse eines anrufenden Anschlußorgans (AO) angesteuert werden, wobei nur die in dem so adressierten Bereich (z. B. TB 1) eingespeicherten Referenzmuster eines einzigen Sprechers oder auch einiger weniger Sprecher dem Vergleichervorgesehen (VG) angeboten werden, und daß eine Umschalteneinrichtung (U) vorgesehen ist, die von einer Steuereinrichtung (ST) so einstellbar ist, daß die über einen Nutzdatenkanal (B-Kanal) ankommende Sprachinformation entweder zum Einspeichern an den Eingang des Referenzmusterspeichers (RMS) gelangt, oder zur Erzeugung von Befehlen (BEF) dem Vergleichervorgesehen (VG) angeboten wird.
2. Spracherkennungseinrichtung nach Anspruch 1, dadurch gekennzeichnet, daß eine Sprachsyntheseeinrichtung (SY) vorgesehen ist, womit ein sprachlicher Dialog zwischen einem anrufenden Teilnehmer und der Spracherkennungseinrichtung (SE) stattfindet, um die Einspeicherung der Referenzmuster in den Referenzmusterspeicher (RMS) für einzeln abzugebende Sprachbefehle eindeutig zu bestimmen.
3. Spracherkennungseinrichtung nach Anspruch 1, dadurch gekennzeichnet, daß für die einem Anschlußorgan (AO) zugehörigen Sprecher vorbestimmte Speicherräume zur Eingabe einer vorbestimmten Anzahl von Befehlen im Referenzmusterspeicher (RMS) zur Verfügung stehen.
4. Spracherkennungseinrichtung nach einem der Ansprüche 2 oder 3, dadurch gekennzeichnet, daß einem Anrufer von der Sprachsyntheseeinrichtung (SY) mitgeteilt wird, wieviele und/oder welche Art von Befehlen noch eingebbar sind.
5. Spracherkennungseinrichtung nach Anspruch 1, dadurch gekennzeichnet, daß der einem Sprecher zugeordnete Speicherraum ganz oder teilweise durch ein von diesem Sprecher einzugebendes Schlüsselwort gegen unbefugten Zugriff geschützt werden kann.
6. Spracherkennungseinrichtung nach Anspruch 5, dadurch gekennzeichnet, daß einzelne Befehle durch ein sprecher-individuelles Schlüsselwort geschützt werden können, so daß andere Sprecher diesen Befehl nicht erzeugen können.
7. Spracherkennungseinrichtung nach Anspruch 1, dadurch gekennzeichnet, daß sprecher-individuell Telefonverzeichnisse eingebbar sind, womit beim späteren Abruf durch sprachliche Eingabe des Namens die Rufnummer der gewünschten Person gewählt wird, wobei die Rufnummer im Referenzmu-

sterspeicher (RMS) jeweils im Zusammenhang mit dem eingegebenen Namen gespeichert wird.

8. Spracherkennungseinrichtung nach Anspruch 7, dadurch gekennzeichnet, daß nach der sprachlichen Eingabe des Namens auch die Rufnummer sprachlich eingegeben wird.

9. Spracherkennungseinrichtung nach Anspruch 7, dadurch gekennzeichnet, daß nach der sprachlichen Eingabe des Namens die Rufnummer mit der Wahl-tastatur eingegeben wird.

Beschreibung

Die Erfindung betrifft eine Spracherkennungseinrichtung für dienste-integrierende Kommunikationsanlagen, insbesondere Fernsprechvermittlungsanlagen nach dem Oberbegriff des Patentanspruches 1.

Aus der DE-OS 31 29 353 ist ein Verfahren zur sprecherunabhängigen Erkennung von einzelnen gesprochenen Worten aus einer vorgegebenen Wortmenge in Fernmeldesystemen bekannt. Es ist ein Referenzmusterspeicher vorgesehen, in den während einer Lernphase durch eine sogenannte Aufbereitungseinheit erzeugte Referenzmuster eingegeben werden, die durch von mehreren Sprechern einmal oder mehrmals sprachlich eingegebenen Worten gebildet werden. Um einen sogenannten Referenzmustersatz zu erstellen, bedarf es einer langen und schwierigen Prozedur, bis die einzelnen als Befehle eindeutig zu erkennenden Worte richtig eingespeichert sind. Beim Abrufen von sprachlich gespeicherten Befehlen ist zwar keine Lernphase mehr erforderlich, jedoch ist die Erkennung relativ unsicher, wenn die Sprecher die zu erkennenden Worte nicht deutlich genug eingeben.

Es ist außerdem ein Verfahren zur sprecherabhängigen Erkennung von einzelnen gesprochenen Worten in Fernmeldesystemen aus der DE-OS 31 29 282 bekannt, wobei ebenfalls ein Sprachmusterspeicher vorgesehen ist. In diesen zentral angeordneten Sprachmusterspeicher werden in Sprachparameterform von einem einzelnen Sprecher abzugebende vorgegebene Worte eingespeichert. Diese Worte sind auf einem Magnetstreifen gespeichert, der sich auf einem Magnetstreifenträger befindet, welcher in ein Eingabegerät eingeschoben wird. Damit wird erreicht, daß der Inhaber einer derartig ausgestalteten Identitätskarte aufgefordert wird, die vorgegebenen Worte in ein Mikrofon zu sprechen. Nachdem die vorher aus dem Magnetstreifen gewonnenen Sprachparameter in den Sprachmusterspeicher des Spracherkennungssystems eingegeben worden sind, kann nun ein Vergleich mit den tatsächlich gesprochenen Worten stattfinden, so daß die Identität einer Person eindeutig erkannt werden kann. Zu diesem Zweck wird die betreffende Person aufgefordert, die in einer Anzeigeeinrichtung schriftlich dargestellten Wort einzusprechen. Ein derartiges Verfahren ist dazu geeignet, das unbefugte Benutzen von Identitätskarten durch fremde Personen auszuschließen. Deshalb ist auch nur ein Kurzzeitspeicher vorgesehen, der jeweils nur die Sprachparameter weniger Worte aufnehmen kann.

Bei einer sprecherabhängigen Erkennung von wahlweise eingebbaren Worten ist pro Sprecher ein relativ großer Speicheraufwand erforderlich, so daß sich sehr umfangreiche Speichereinrichtungen ergeben, wenn für viele Sprecher eine jeweils sprecherabhängige Spracherkennung vorgesehen wird. Die einzelnen jeweils einem Sprecher zugeordneten Bereiche müssen dann beispielsweise durch Eingabe von Kennziffern individuell

adressiert werden. Es ist dann außerdem eine relativ umständliche Lernphase erforderlich, um eine hohe Erkennungssicherheit zu erreichen.

Die Aufgabe der Erfindung besteht darin, eine Spracherkennungseinrichtung für dienste-integrierende Kommunikationsanlagen, insbesondere Fernsprechvermittlungsanlagen vorzuschlagen, wobei jeweils die Vorteile von sprecherunabhängigen und sprecherabhängigen Spracherkennungsverfahren genutzt werden, ohne daß die bei dem einen oder bei dem anderen Verfahren vorhandenen Nachteile auftreten.

Diese Aufgabe wird durch eine Merkmalskombination gelöst, wie sie im Patentanspruch 1 angegeben ist.

Damit wird in vorteilhafter Weise erreicht, daß sowohl eine sprecherabhängige als auch eine sprecherunabhängige Spracherkennung vorgesehen sein kann. Wenn in dem einem Anschlußorgan zugeordneten Bereich des Referenzmusterspeichers die Referenzmuster nur eines einzigen Sprechers eingespeichert sind, so liegt eine sprecherabhängige Spracherkennung vor, wobei automatisch eine direkte Adressierung durch die Kennung dieses Anschlußorgans gegeben ist. Eine sprecherunabhängige Spracherkennung liegt dann vor, wenn in den einem Anschlußorgan zugeordnete Bereiche die Sprachmuster mehrerer Sprecher eingegeben sind. Dabei ist jedoch die Anzahl der Sprecher auf einige wenige begrenzt, so daß auch dabei die Vorteile der sprecherabhängigen Spracherkennung erhalten bleiben, weil die Sprachmuster dieser wenigen Sprecher nacheinander mit dem eingesprochenen Wort verglichen werden können.

In den Unteransprüchen werden vorteilhafte Weiterbildungen der Erfindung angegeben. Durch eine zusätzliche Sprachsyntheseeinrichtung kann, ein sprachlicher Dialog zwischen einem anrufenden Teilnehmer und der Spracherkennungseinrichtung stattfinden, damit dem Teilnehmer vorgegeben werden kann, wieviele und oder welche Arten von Befehlen eingegbar sind. Es wird dabei außerdem vom Teilnehmer erkannt, ob das Sprachmuster für einen einzugebenden Befehl richtig abgespeichert ist. Der durch die Kennung eines Anschlußorgans automatisch angesteuerte Bereich im Referenzmusterspeicher kann außerdem einem Sprecher individuell, zugeordnete Speicherräume enthalten, die durch ein einfaches Kennzeichen oder ein Schlüsselwort von dem betreffenden Sprecher aus direkt ansteuerbar sind.

Ein Ausführungsbeispiel der Erfindung wird nachfolgend anhand von Zeichnungen näher erläutert. Es zeigt

Fig. 1 das Blockschaltbild einer dienste-integrierenden Kommunikationsanlage,

Fig. 2 das Blockschaltbild einer Spracherkennungseinrichtung.

In der Fig. 1 ist dargestellt, wie Fernsprechgeräte *FE 1*, *FE 2* über Anschlußorgane *AO 1*, *AO 2* an eine Vermittlungsanlage *VA* angeschlossen sind. Neben anderen nicht dargestellten Einrichtungen kann an eine derartige Vermittlungsanlage *VA* eine Spracherkennungseinrichtung *SE* angeschlossen sein, wobei der Austausch von Nutzdaten über einen sogenannten B-Kanal stattfindet. Es ist dabei wie auch für die Anschlußorgane *AO* ein Signalisierungskanal (D-Kanal) vorgesehen, worüber neben den Nutzdaten zusätzliche vermittlungstechnische Informationen übertragen werden können. Über diesen D-Kanal erhält die Spracherkennungseinrichtung *SE* die Adresse und/oder die Rufnummer eines anrufenden Anschlußorgans, z. B. *AO 1*. Die Spracherkennungseinrichtung *SE* erkennt entspre-

chend den im Referenzmusterspeicher vorgegebenen Sprachmustern bestimmte über den Nutzdatenkanal (B-Kanal) einlaufende sprachlich abgegebene Worte. Bei eindeutiger Erkennung werden daraus Befehle erzeugt, die an eine Datenverarbeitungsanlage *DVA* weitergegeben werden können, wobei es sich beispielsweise um ein Auskunftssystem oder um einen zentralen Sprachaufnahmespeicher (Voice-Mail-System) handeln kann. Mit den von der Spracherkennungseinrichtung *SE* erzeugten Befehlen können jedoch auch vermittlungstechnische Prozeduren innerhalb der Vermittlungsanlage *VA* ausgelöst werden.

Das in Fig. 2 dargestellte Blockschaltbild zeigt die einzelnen Komponenten der Spracherkennungseinrichtung *SE*. Um die Spracherkennung zu ermöglichen, müssen zunächst Sprachmuster für die Worte gebildet werden, die bei einer sprachlichen Eingabe von Befehlen erkannt werden sollen. Dies geschieht dadurch, daß bei Ansteuerung der Spracherkennungseinrichtung *SE* die über den Nutzdatenkanal, B-Kanal, einlaufende Sprachinformation in einer Vorverarbeitungseinheit *VVE* aufgenommen wird. Dort werden Referenzmuster von jedem gesprochenen Wort gebildet, welche über die im Ruhezustand befindliche Umschalteneinrichtung *U* an den Eingang des Referenzmusterspeichers *RMS* gelangen. Da über den Signalisierungskanal, D-Kanal, gleichzeitig die Kennzeichnung eines anrufenden Anschlußorgans *AO* mitgeteilt wird, kann die D-Kanal-Steuerung *DS* daraus eine Adresse errechnen. Mit dieser Adresse wird ein so bestimmter Teilbereich *TB* im Referenzmusterspeicher *RMS* angesteuert. Für einen Anrufer besteht nun die Möglichkeit, in diesen Teilbereich, z. B. *TB 1*, mehrere Befehlsworte nacheinander einzugeben.

Mit Hilfe einer zusätzlich vorgesehenen Sprachsyntheseeinrichtung *SY* kann dem anrufenden Teilnehmer mitgeteilt werden, daß die von ihm gemachte Spracheingabe richtig verstanden worden ist. Zu diesem Zweck schaltet die Steuereinrichtung *ST* nach dem Einspeichern des Referenzmusters eines Befehlswortes den Referenzmusterspeicher *RMS* auf Lesen, so daß mit Hilfe des Referenzmusters ein Quittungswort über die Sprachsyntheseeinrichtung *SY* und den Nutzdatenkanal, B-Kanal, zum Anschlußorgan des anrufenden Teilnehmers gelangt. Danach kann der anrufende Teilnehmer erkennen, ob das von ihm eingegebene Befehlswort richtig erkannt worden ist.

Auf die gleiche Weise lassen sich auch gleichartige Befehlsworte von verschiedenen Sprechern in den Referenzmusterspeicher *RMS* eingeben, die in den gleichen Teilbereich, z. B. *TB 1* gelangen, wenn diese verschiedenen Sprecher vom gleichen Anruforgan, z. B. *AO 1*, aus anrufen. Innerhalb eines Teilbereichs, z. B. *TB 1*, können dabei für jeden Sprecher eigene Speicherräume vorgesehen sein. Wenn der Referenzmusterspeicher *RMS* so aufgeteilt ist, daß in jedem der einzelnen Teilbereiche *TB 1* bis *TB n* sprecherindividuelle Speicherräume zur Verfügung stehen, so muß ein spezielles Kennzeichen abgegeben werden, damit ein solcher Speicherraum gezielt adressiert werden kann.

Es ist jedoch auch möglich, gleichartige Befehlsworte von verschiedenen Sprechern aus einzugeben, ohne daß der betreffende Teilbereich, z. B. *TB 1*, unterteilt ist. Gleichartige Befehle verschiedener Sprecher werden dann nebeneinander abgelegt, so daß beim Abrufen eines Befehls die betreffenden Sprachreferenzmuster nacheinander überprüft werden können.

Wenn auf diese Weise die zu Befehlen führenden

Sprachreferenzmuster von den beteiligten Sprechern eingegeben worden sind, so wird beim erneuten Anrufen der Spracherkennungseinrichtung *SE* aus jedem eingegebenen Befehlswort ein Befehl *BEF* erzeugt, wenn eine Übereinstimmung mit dem im Referenzmusterspeicher *RMS* eingespeicherten Sprachreferenzmusterwort festgestellt wird. Dies geschieht dadurch, daß durch die Steuereinrichtung *ST* die Umschalteneinrichtung *U* in ihre Arbeitslage versetzt wird. Bei dieser Schaltstellung der Umschalteneinrichtung *U* gelangen die von der Vorverarbeitungseinrichtung *VVE* errechneten Referenzmuster aus den über den Nutzdatenkanal, B-Kanal, einlaufenden Sprachsignalen nicht zum Referenzmusterspeicher *RMS* sondern zu einem Vergleich *VG*. Gleichzeitig wird der Referenzmusterspeicher *RMS* von der Steuereinrichtung *ST* auf Lesen geschaltet, so daß der Ausgang des Referenzmusterspeichers *RMS* mit dem anderen Eingang des Vergleichers *VG* verbunden ist. Da über den Signalisierungskanal, D-Kanal, die Identität des anrufenden Anschlußorgans, z. B. *AO 1*, empfangen worden ist, wird der betreffende Teilbereich, z. B. *TB 1*, angesteuert. Wenn vom Vergleich *VG* eine Übereinstimmung eines angekommenen Sprachreferenzmusters mit einem im Referenzmusterspeicher *RMS* eingespeicherten Referenzmuster festgestellt wird, so wird dieses an die Steuereinrichtung *ST* gemeldet. Da der Steuereinrichtung *ST* gleichzeitig bekannt ist, an welcher Stelle des Referenzmusterspeichers *RMS* eine Übereinstimmung gefunden wurde, kann daraus unmittelbar der richtige Befehl *BEF* abgeleitet und ausgegeben werden.

Im Referenzmusterspeicher *RMS* können neben den eigentlichen Sprachreferenzmustern auch andere Informationen abgelegt sein, wobei es sich beispielsweise um Rufnummern handeln kann. Diese Rufnummern werden dann ausgegeben, wenn bei einem sprachlich eingegebenen Befehlsabruf der Name eines anzurufenden Teilnehmers erkannt wird, wovon das entsprechende Sprachreferenzmuster im Referenzmusterspeicher *RMS* abgelegt ist. Die Eingabe der für eine derartige Funktion bereitzustellenden Rufnummer erfolgt dadurch, daß zunächst wie vorher beschrieben worden ist, der Name eines anzurufenden Teilnehmers eingegeben wird, nach dem zuvor ein besonderes Kennzeichen für die Einspeicherung von sprachlich abrufbaren Verbindungswünschen abgegeben wurde. Danach kann die zugehörige Rufnummer in Einzelziffern entweder sprachlich oder mit Hilfe der Wähltastatur eingegeben werden, so daß diese im Referenzmusterspeicher *RMS* oder in einem parallel adressierbaren Zusatzspeicher zugeordnet zum Namen des anzurufenden Teilnehmers abgelegt werden kann. Bei einer derartigen Prozedur kann dem anrufenden Teilnehmer mit Hilfe der Sprachsyntheseeinrichtung *SY* Hilfestellung gegeben werden, indem er zur richtigen Zeit aufgefordert wird, die einzelnen Ziffern einer Rufnummer einzusprechen oder einzutasten.

Wenn auf diese Weise sprecher-individuelle Telefonverzeichnisse in kleinem Umfang gebildet werden können, so kann außerdem für die Adressierung eines solchen Bereiches im Referenzmusterspeicher *RMS* ein Paßwort vorgegeben werden, so daß ein anderer Sprecher keinen Zugang zu einem solchen Speicherraum bekommt.

Bei der Organisation des Referenzmusterspeichers *RMS* sind viele Variationen möglich, so daß die Spracherkennungseinrichtung *SE* auf den jeweiligen Anwendungsfall optimal angepaßt werden kann. Es ist dabei

von Vorteil, daß die einzelnen Teilbereiche *TB 1* bis *TBn* jeweils vorbestimmten Anschlußorganen *AO 1*, *AO 2* usw. fest zugeordnet sind. Diese Zuordnung kann in der inneren D-Kanal-Steuerung *DS* programmiert werden. Damit ist auch gleichzeitig festgelegt, daß nur von solchen Anschlußorganen *AO* aus Spracheingaben in die Spracherkennungseinrichtung *SE* möglich sind, für die ein Teilbereich *TB* im Referenzmusterspeicher *RMS* eröffnet worden ist.

3643838

12 14
Nummer: 36 43 838
Int. Cl. 4: G 10 L 3/00
Anmeldetag: 20. Dezember 1986
Offenlegungstag: 30. Juni 1988

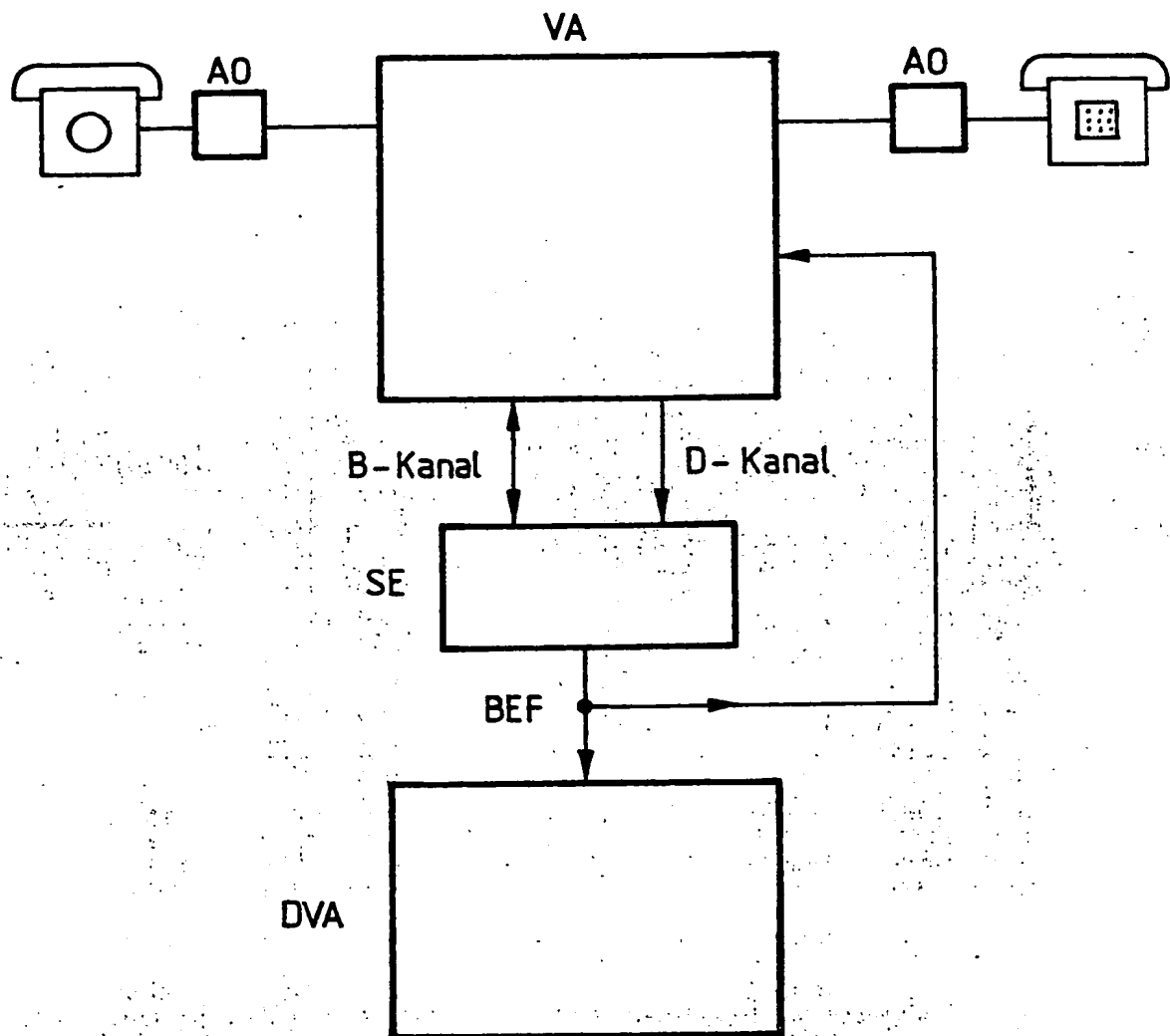


Fig. 1

3643838

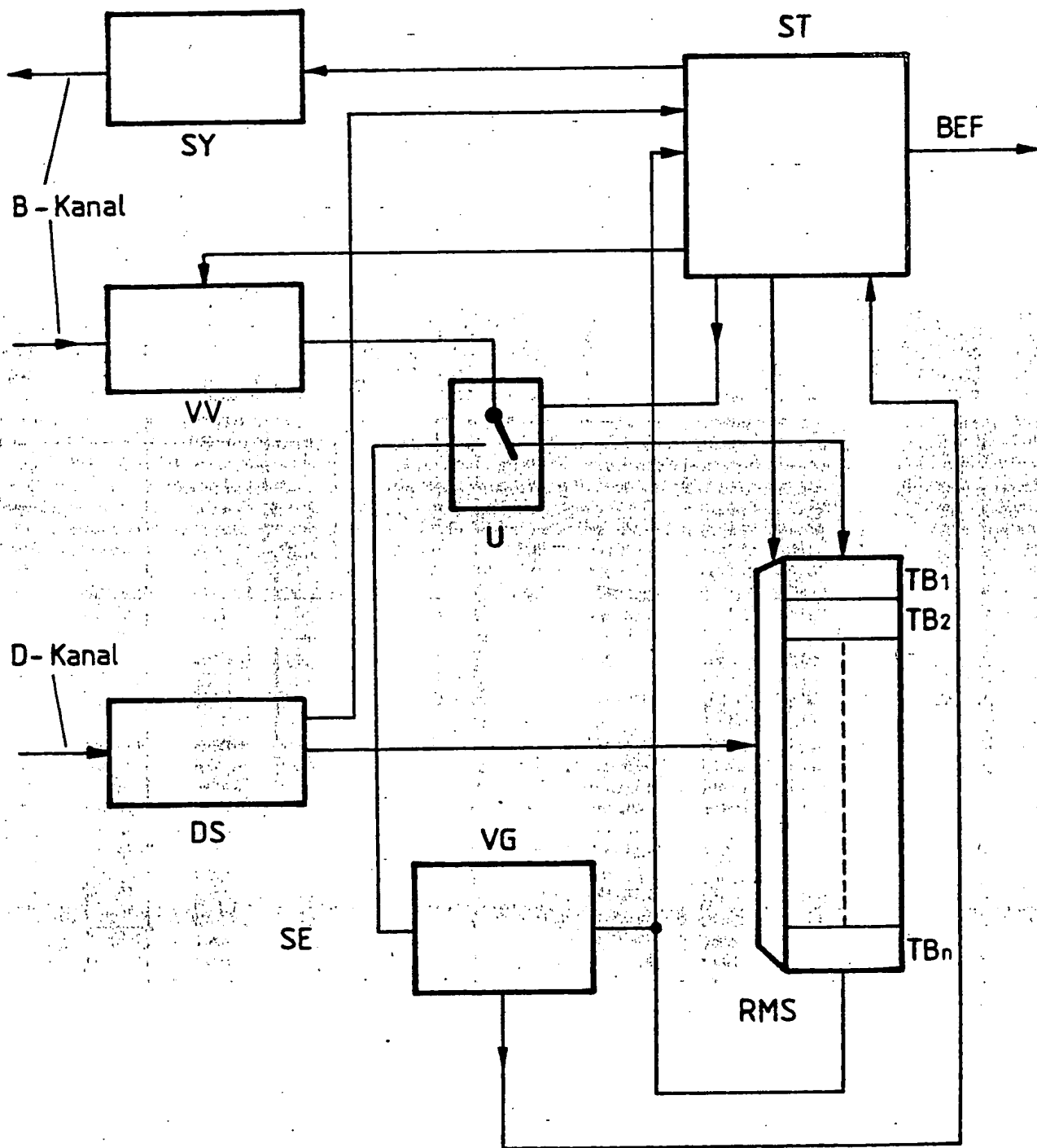


Fig. 2